[Submitted on 27 Aug 2024 (v1), last revised 11 Feb 2025 (this version, v2)]

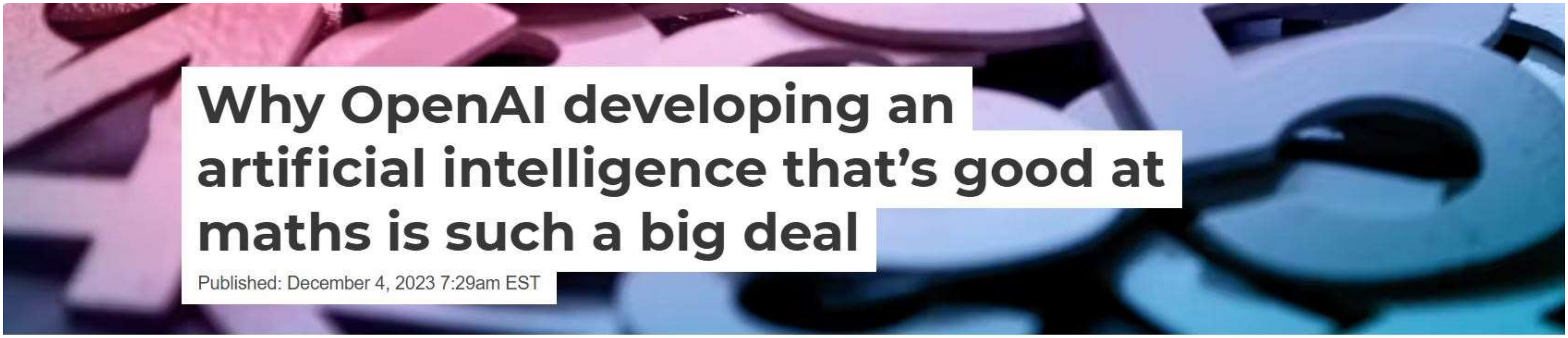
What makes math problems hard for reinforcement learning: a case study

Ali Shehper, Anibal M. Medina-Mardones, Lucas Fagan, Bartłomiej Lewandowski, Angus Gruen, Yang Qiu, Piotr Kucharski, Zhenghan Wang, Sergei Gukov

Using a long-standing conjecture from combinatorial group theory, we explore, from multiple perspectives, the challenges of finding rare instances carrying disproportionately high rewards. Based on lessons learned in the context defined by the Andrews-Curtis conjecture, we propose algorithmic enhancements and a topological hardness measure with implications for a broad class of search problems. As part of our study, we also address several open mathematical questions. Notably, we demonstrate the length reducibility of all but two presentations in the Akbulut-Kirby series (1981), and resolve various potential counterexamples in the Miller-Schupp series (1991), including three infinite subfamilies.

Comments: 58 pages, 25 figures, 1 table. Try it: this https URL

Subjects: **Machine Learning (cs.LG)**; Artificial Intelligence (cs.AI); Combinatorics (math.CO); Group Theory (math.GR); Geometric Topology (math.GT)



Why OpenAI developing an artificial intelligence that's good at maths is such a big deal

Published: December 4, 2023 7:29am EST

Daboost / Shutterstock

Email
X (Twitter)
Facebook 67
LinkedIn
Print

With the recent sacking and swift rehiring of Sam Altman by OpenAI, debates around the development and use of artificial intelligence (AI) are once again in the spotlight. What's more unusual is that a prominent theme in media reporting has been the ability of AI systems to do maths.

related to the
called Q*. The
vance and one of
athematically.

Author

Tom Oliver
Lecturer, Computer Science and Engineering, University of Westminster

Disclosure statement

Tom Oliver does not work for, consult, own shares in or receive funding from any company or organization that would benefit from this article, and has disclosed no relevant affiliations beyond their academic appointment.

OpenAI Scale Ranks Progress Toward 'Human-Level' Problem Solving

The company believes its technology is approaching the second level of five on the path to artificial general intelligence

Sam Altman, CEO of OpenAI

Dario Amodei

Machines of Loving Grace[1]

How AI Could Transform the World for the Better

October 2024

I think and talk a lot about the risks of powerful AI. The company I'm the CEO of, Anthropic, does a lot of research on how to reduce these risks. Because of this, people sometimes draw the conclusion that I'm a pessimist or "doomer" who thinks AI will be mostly bad or dangerous. I don't think that at all. In fact, one of my main reasons for focusing on risks is that they're the only thing standing between us and what I see as a fundamentally positive future. **I think that most people are**

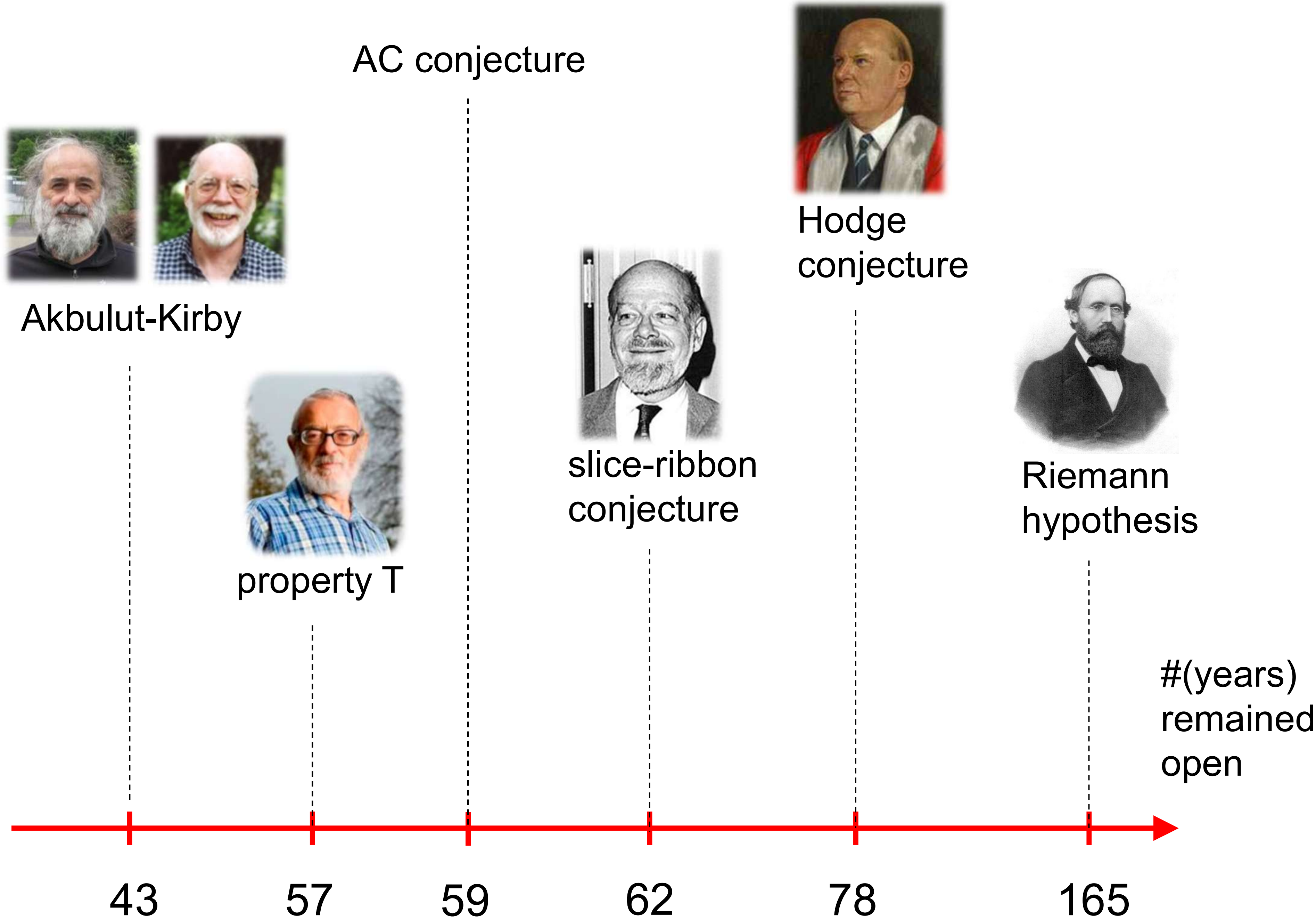
- It's smart: "It is smarter than a Nobel Prize winner across most relevant fields."
- It's multimodal: It isn't limited to one mode of action. It can interface through text, audio, video, mouse and keyboard control, and the internet.
- It's independent: "It does not just passively answer questions," he wrote. "Instead, it can be given tasks that take hours, days, or weeks to complete, and then goes off and does those tasks autonomously."
- It's abstract: "It does not have a physical embodiment."
- It's replicable: "The resources used to train the model can be repurposed to run millions of instances of it."
- It's fast. "The model can absorb information and generate actions at roughly 10x-100x human speed."
- It's cooperative. "Each of these million copies can act independently on unrelated tasks, or if needed can all work together in the same way humans would collaborate."
- Amodei thinks this form of AI might arrive as early as 2026. In a nutshell, he describes it as "a country of geniuses in a datacenter."

Mathematical problems range in complexity and difficulty:

- elementary school problems
- middle school problems
- high school problems

general purpose tool

- math olympiad problems; e.g. International Mathematical Olympiad (IMO)
- undergraduate-level problems
- graduate-level problems
- simple research problems that mathematicians solve as part of their daily routine
- research problems whose solutions become the main results of research papers published in leading mathematical journals
- Long-standing problems open for decades (AC, …)
- Millennium Prize Problems


AC conjecture
Akbulut-Kirby
property T
slice-ribbon conjecture
Hodge conjecture
Riemann hypothesis
#(years) remained open
43
57
59
62
78
165

1 out of 2.46×10^{26} colorings of a complete graph on 17 nodes

Values / known bounding ranges for Ramsey numbers $R(r, s)$ (sequence A212954 in the OEIS)

r \ s	1	2	3	4	5	6	7	8	9	10
1	**1**	1	1	1	1	1	1	1	1	1
2		**2**	3	4	5	6	7	8	9	10
3			**6**	9	14	18	23	28	36	40–41[14]
4				**18**	25[9]	36–40	49–58	59[15]–79	73–105	92–135
5					**43–46**[11]	59[16]–85	80–133	101–193	133–282	149[15]–381
6						**102–160**	115[15]–270	134[15]–423	183–651	204–944
7							**205–492**	219–832	252–1368	292–2119
8								**282–1518**	329–2662	343–4402
9									**565–4956**	581–8675
10										**798–16064**

63. Cantor's Theorem

Author: mathlib

```
theorem Cardinal.cantor (a : Cardinal.{u}) :
    a < 2 ^ a
```

docs, source

64. L'Hopital's Rule

Author: Anatole Dedecker

```
theorem deriv.lhopital_zero_nhds {a : ℝ} {l : Filter ℝ} {f : ℝ → ℝ} {g : ℝ → ℝ}
        (hdf : ∀ᶠ (x : ℝ) in nhds a, DifferentiableAt ℝ f x)
        (hg' : ∀ᶠ (x : ℝ) in nhds a, deriv g x ≠ 0)
        (hfa : Filter.Tendsto f (nhds a) (nhds 0)) (hga : Filter.Tendsto g (nhds a) (nhds 0))
        (hdiv : Filter.Tendsto (fun (x : ℝ) => deriv f x / deriv g x) (nhds a) l) :
    Filter.Tendsto (fun (x : ℝ) => f x / g x) (nhdsWithin a {a}ᶜ) l
```

docs, source

65. Isosceles Triangle Theorem

Author: Joseph Myers

Source: https://leanprover-community.github.io/100.html

1 Mathematics = Game

2 Hard Math = Long Game

HRL
multiscale

3 What makes math problems hard?

- for a human = - for an AI model

Conjecture [J.Andrews and M.Curtis '65]:

Every balanced presentation of the trivial group

$$\langle\, x_1, \ldots, x_n \mid r_1, \ldots, r_n \,\rangle$$

can be reduced to the trivial presentation

$$\langle\, x_1, \ldots, x_n \mid x_1, \ldots, x_n \,\rangle$$

by a sequence of Andrews-Curtis (Nielsen) moves:

$$r_i,\, r_j \;\mapsto\; r_i r_j,\, r_j$$

$$r_i \;\mapsto\; r_i^{-1}$$

"handle slides"

$$r_i \;\mapsto\; x_j^{\pm 1} r_i x_j^{\mp 1}$$

$$^{*}\langle\, x_1, \ldots, x_n \mid r_1, \ldots, r_n \,\rangle \;\leftrightarrow$$

"handle cancellation"

$$\leftrightarrow\; \langle\, x_1, \ldots, x_n, x_{n+1} \mid r_1, \ldots, r_n, x_{n+1} \,\rangle$$

---

* generalized

THE COMPLEXITY OF BALANCED PRESENTATIONS AND THE ANDREWS–CURTIS CONJECTURE

MARTIN R. BRIDSON

Hard AC presentations

arXiv:1504.04187

Theorem A. *For $k \geq 4$ one can construct explicit sequences of k-generator balanced presentations $\mathcal{P}_n$ of the trivial group so that*

(1) the presentations $\mathcal{P}_n$ are AC-trivialisable;
(2) the sum of the lengths of the relators in $\mathcal{P}_n$ is at most $24(n+1)$;
(3) the number of (dihedral) AC moves required to trivialise $\mathcal{P}_n$ is bounded below by the function $\Delta(\lfloor \log_2 n \rfloor)$ where $\Delta : \mathbb{N} \to \mathbb{N}$ is defined recursively by $\Delta(0) = 2$ and $\Delta(m+1) = 2^{\Delta(m)}$.

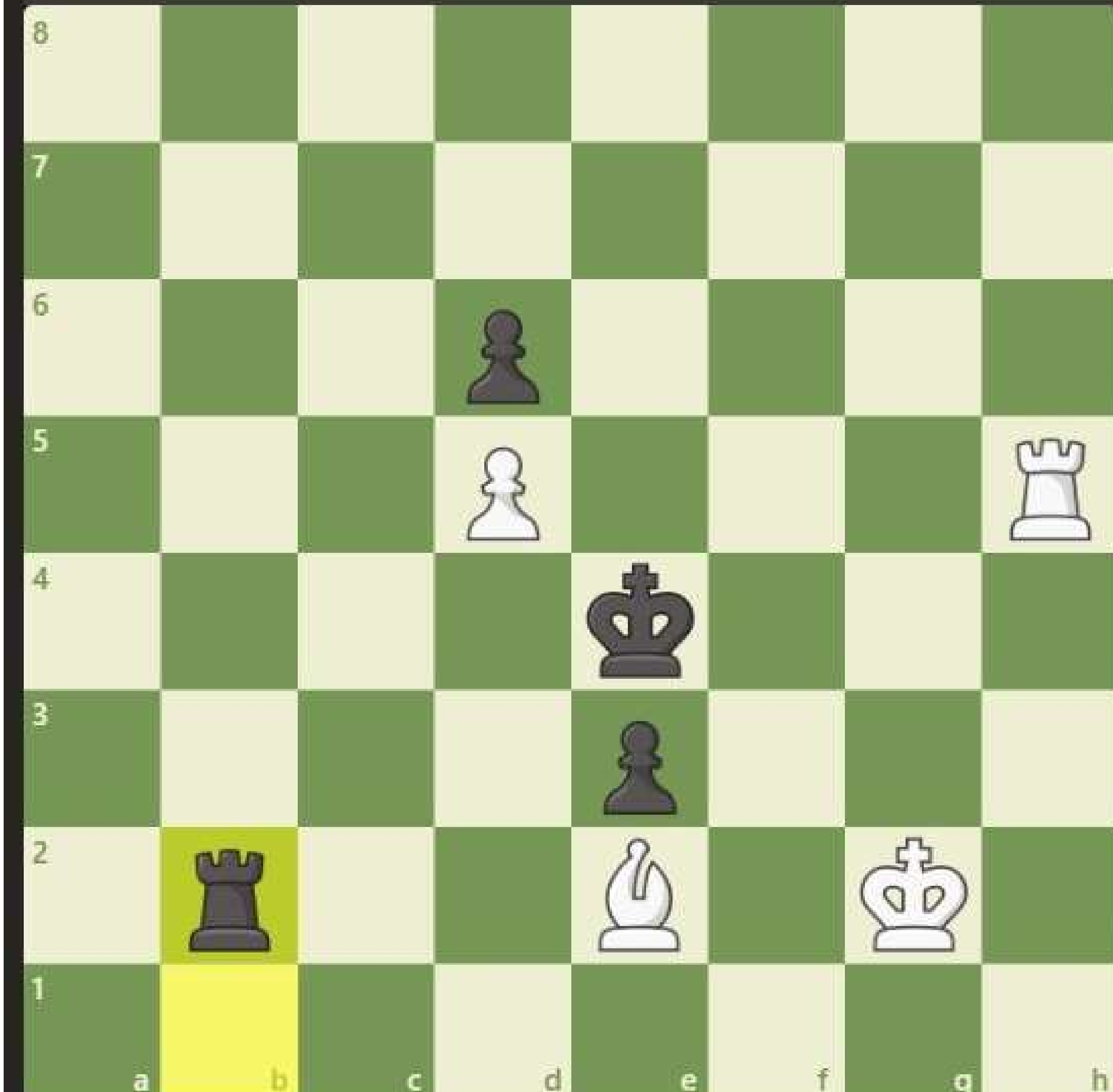
7.4. **An Example.** Let me close by writing down an explicit presentation to emphasize that the explosive growth in the length of AC-trivialisations begins with relatively small presentations. Here is a balanced presentation of the trivial group that requires more than 10^{10000} AC-moves to trivialise it. We use the commutator convention $[x, y] = xyx^{-1}y^{-1}$.

$$\langle a, t, \alpha, \tau \mid [tat^{-1}, a]a^{-1}, \quad [\tau\alpha\tau^{-1}, \alpha]\alpha^{-1}, \quad \alpha t^{-1}\alpha^{-1}[a, [t[t[ta^{20}t^{-1}, a]t^{-1}, a]t^{-1}, a]], \quad a\tau^{-1}a^{-1}[\alpha, [\tau[\tau[\tau\alpha^{20}\tau^{-1}, \alpha]\tau^{-1}, \alpha]\tau^{-1}, \alpha]]\rangle.$$

The game lasted over 20 hours to end in a draw!

Ivan Nikolic vs. **Goran Arsovic**

1/2-1/2 | Belgrade / Belgrade YUG / 17 Feb 1989 | ECO: E95



259.	Bc4	Rg3
260.	Rb2	Rg5+
261.	Bd5	Rg3
262.	Rh2	Rc3+
263.	Bc4	Rg3
264.	Rh8	Ka3
265.	Ra8+	Kb2
266.	Ra2+	Kb1
267.	Rf2	Kc1
268.	Kd4	Kd1
269.	Bd3	Rg7
1/2-1/2		

initial state

$\underbrace{\boxed{\sigma_1}\boxed{\sigma_2}\boxed{}\cdots\boxed{}}_{L}$

sequence of operations / actions

goal state

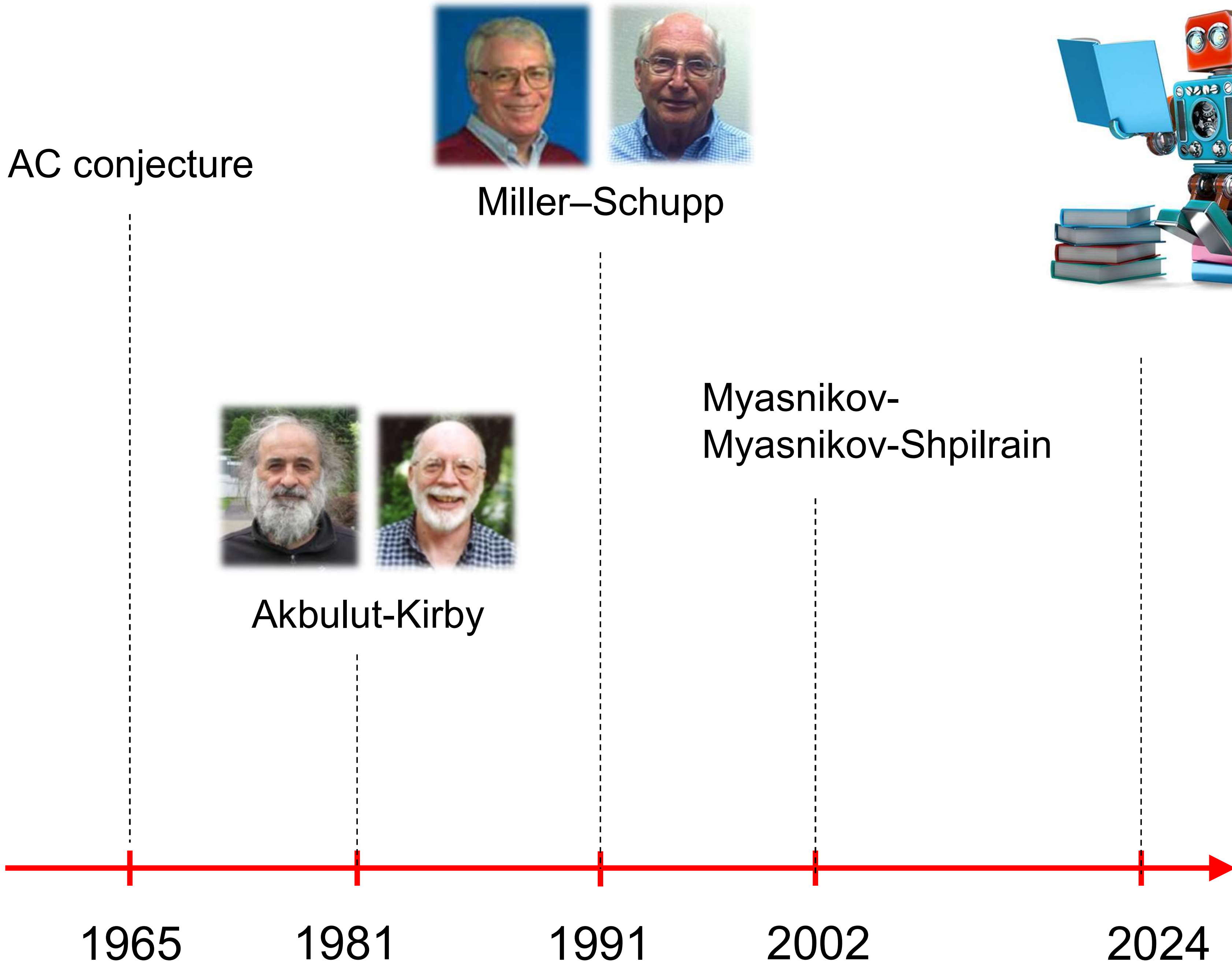
$\sim L^n$ easy

- or -

initial state $'$

$\boxed{\sigma_1'}\boxed{\sigma_2'}\boxed{}\cdots\boxed{}$

$\sim \underbrace{2^{2^{2^{2^{\cdot^{\cdot^{\cdot^{2^{2}}}}}}}}}_{L \text{ times}}$ hard

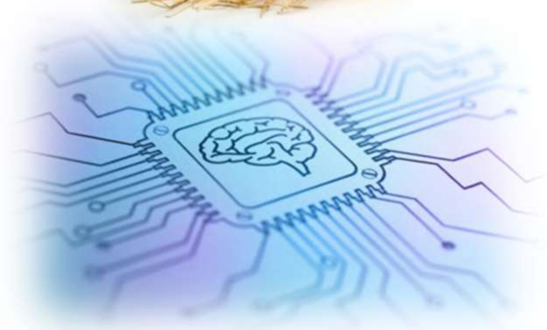

AC conjecture
Miller–Schupp
Akbulut-Kirby
Myasnikov-
Myasnikov-Shpilrain
1965
1981
1991
2002
2024

$$\mathrm{AK}(n) = \langle x, y \mid x^n = y^{n+1}, xyx = yxy \rangle, \quad n \geq 3$$

Theorem 1. *For every $n \geq 2$, $AK(n)$ is AC-equivalent to the presentation*

$$\langle x, y \mid x^{-1}yx = xyx^{-1}y \ , \ xyx = yx^{n-1}y \rangle,$$

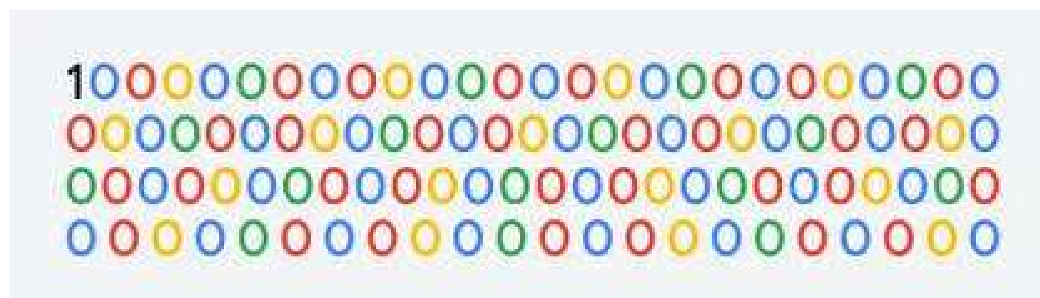
of length $n + 11$. This gives a reduction in length of $AK(n)$ for all $n \geq 5$.



Length of RL path = 1,546 steps

12 basic moves

$12^{1546} \sim 10^{1668}$ search size



S.Ali, A.Medina-Mardones, B.Lewandowski, A.Gruen, P.Kucharski, L.Fagan, Y.Qiu, Z.Wang, S.G

- Solved 700+ Miller–Schupp presentations

1 Mathematics = Game

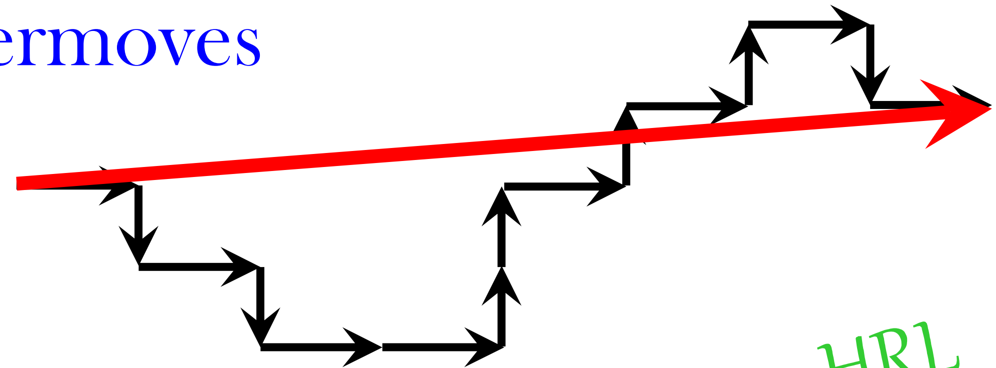
2 Hard Math = Long Game

HRL
multiscale

3 Solution:

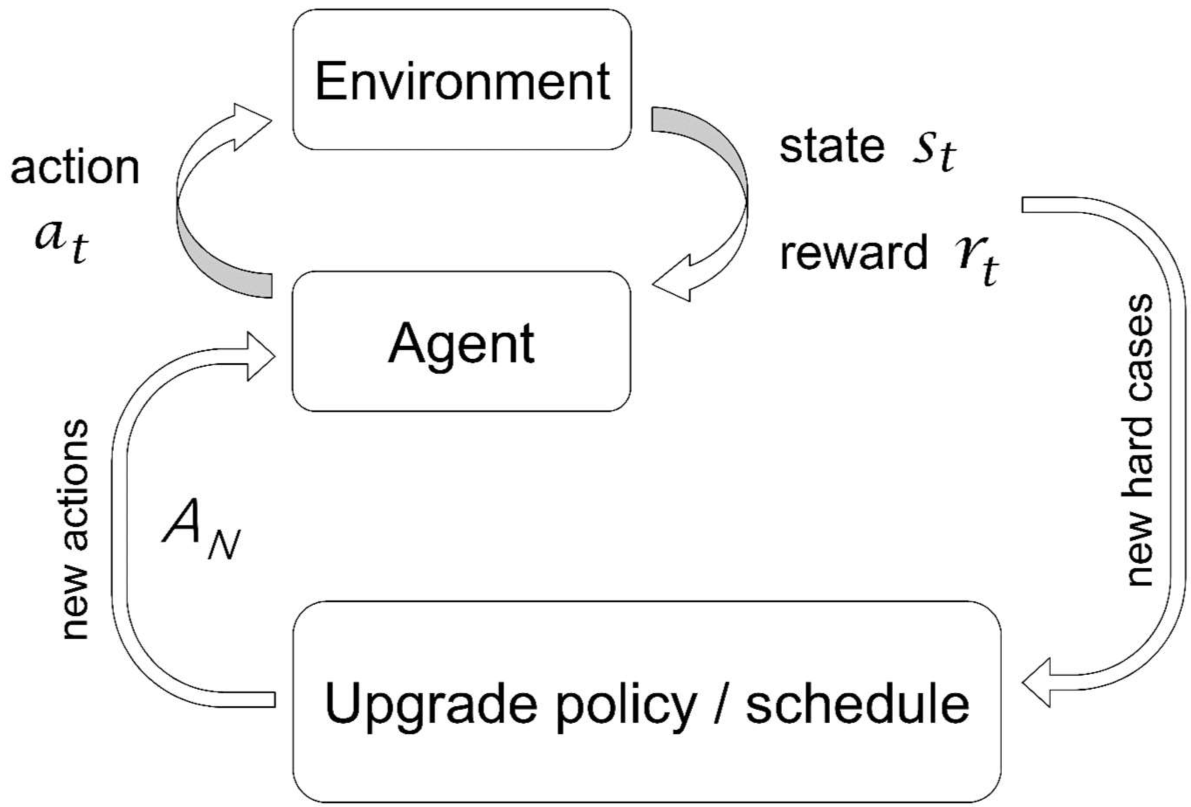
- add and remove Supermoves
- adaptive / dynamic action space (updates)

- add and remove Supermoves

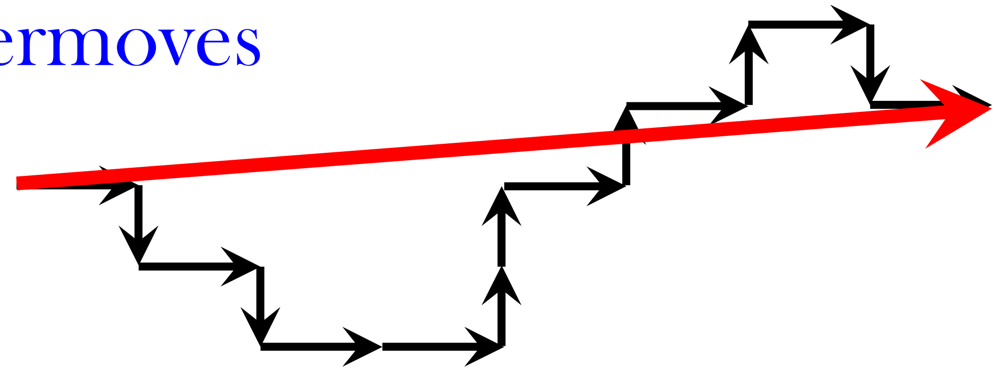


HRL
multiscale

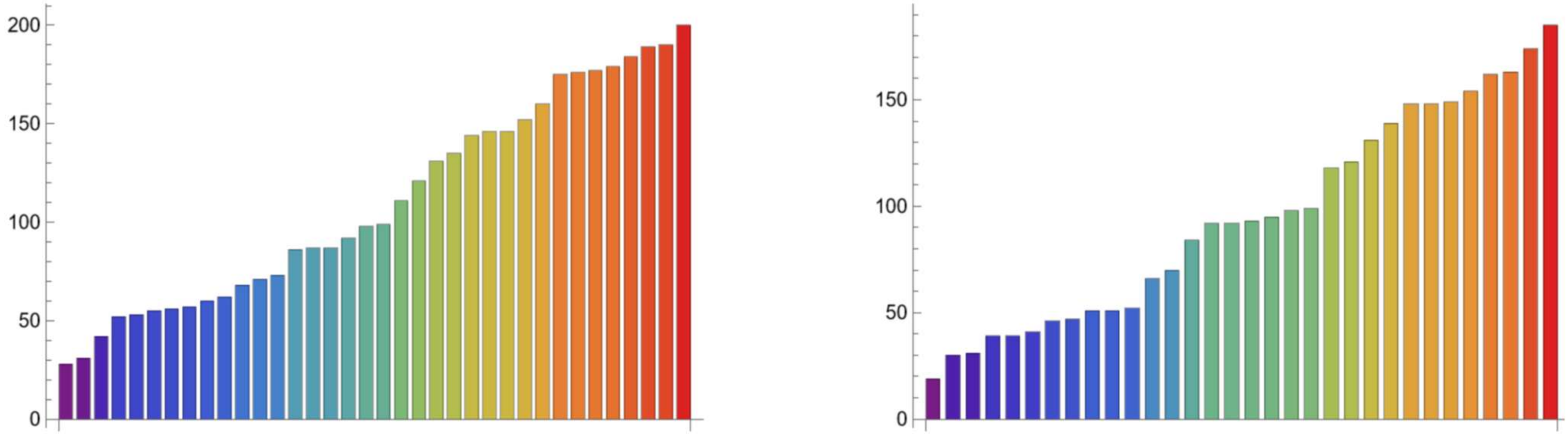
- adaptive / dynamic action space: updates



- add and remove Supermoves



- adaptive / dynamic action space: updates

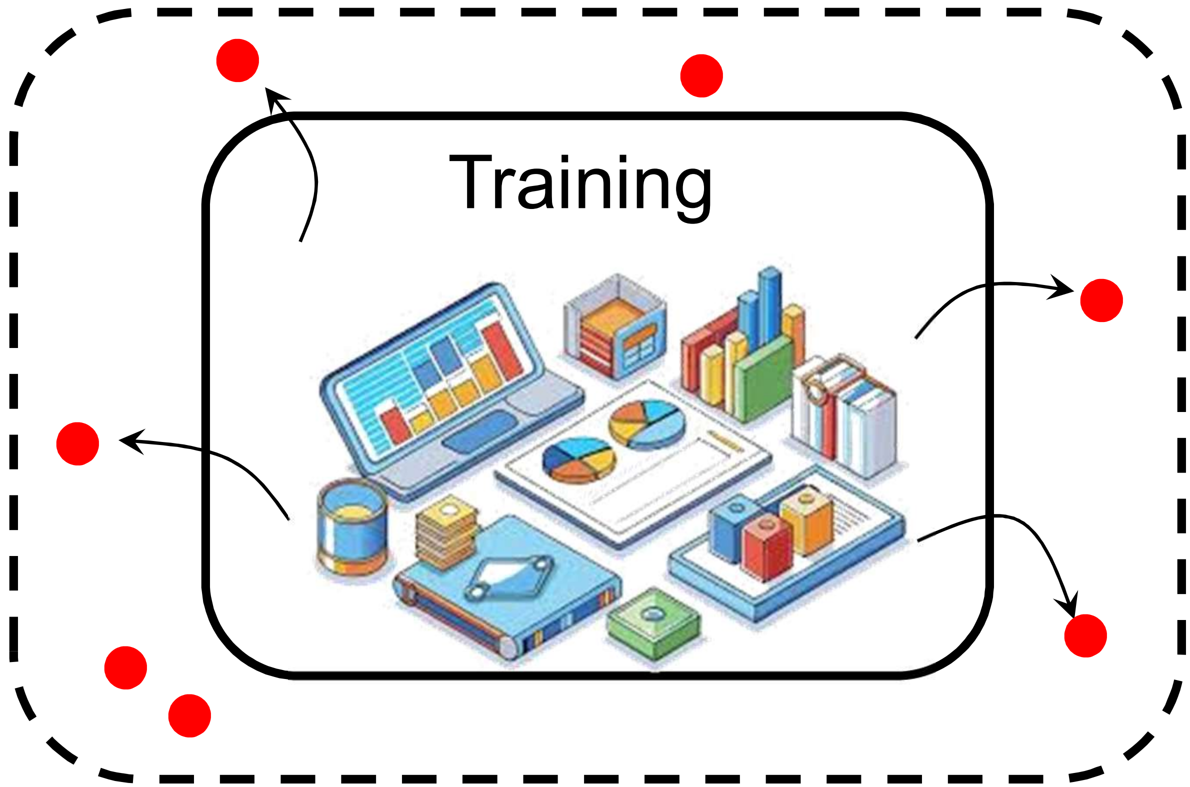


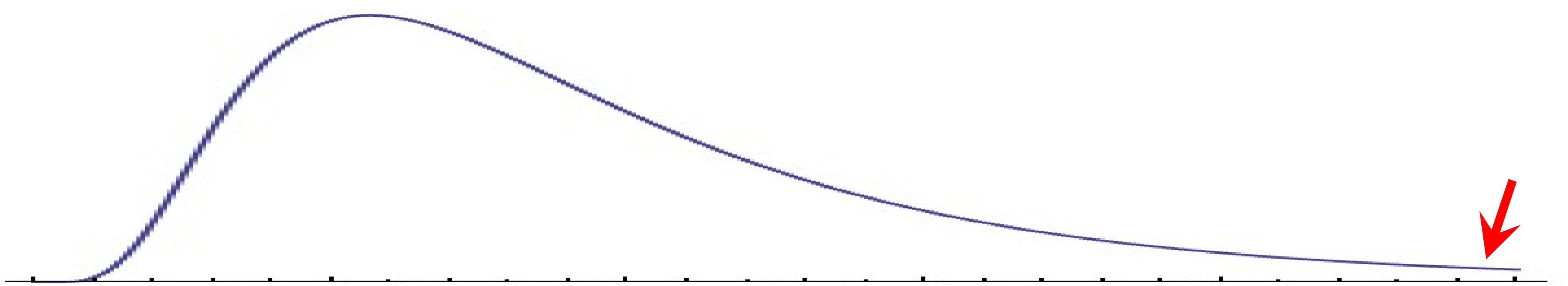
(A) all

(B) GS-solved

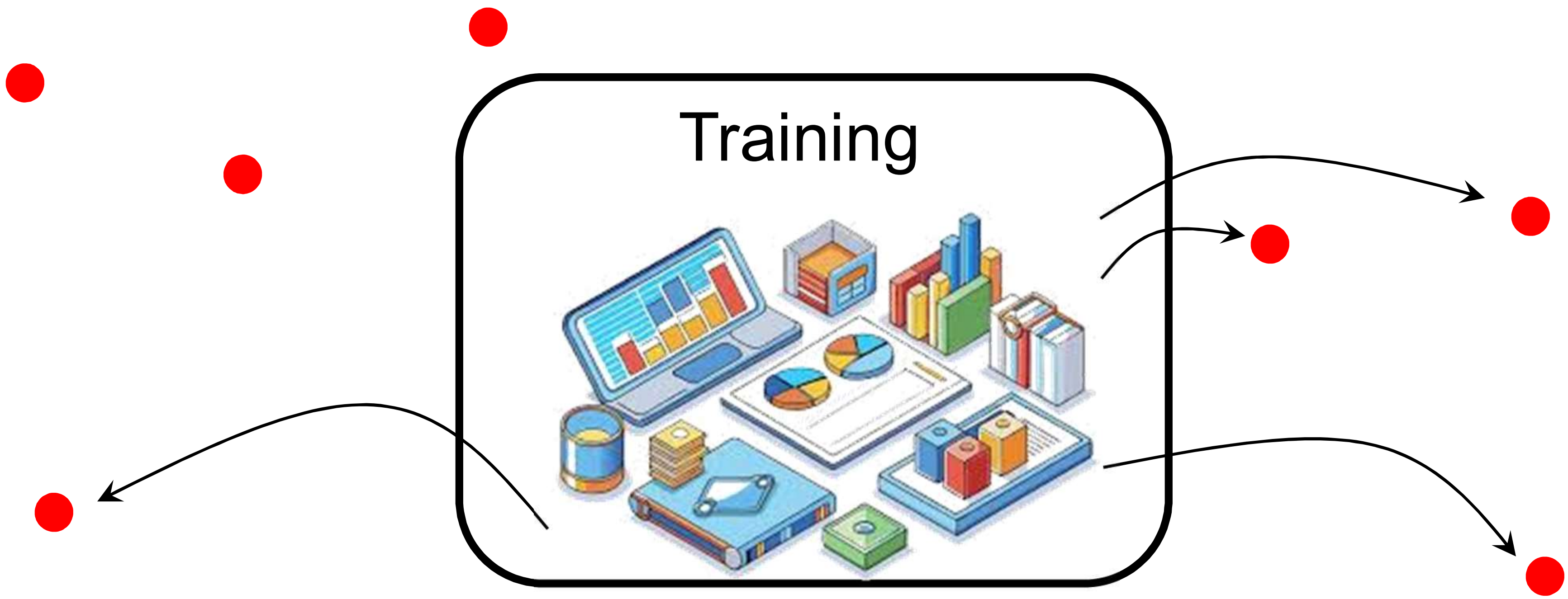
FIGURE 8. Path length distributions for AC presentations solved by the RL agent at $N = 8 \times 10^7$ from *all* and *GS-solved* datasets are nearly identical. In both cases, hard instances are shown in red.

- generalization, OOD performance, robustness

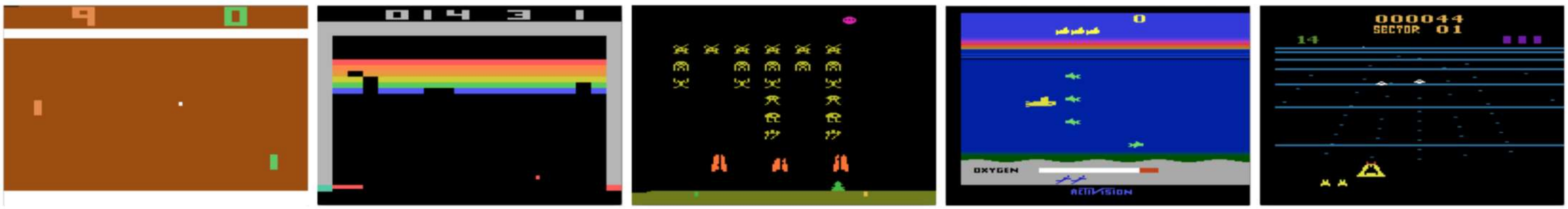




- generalization, OOD performance, robustness



- very different from games (problems) where agent can learn to fail quickly



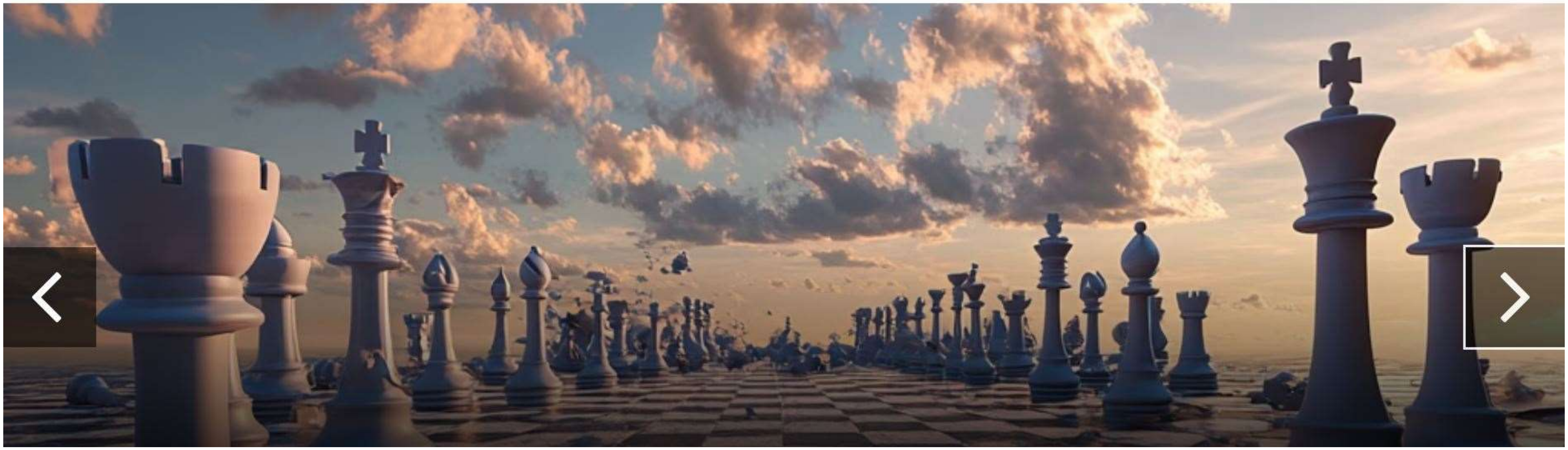
Vladimir I. Arnold

“Mathematics is the part of physics where experiments are cheap.”

https://math-ai.caltech.edu

Caltech | Math + AI = AGI

Menu



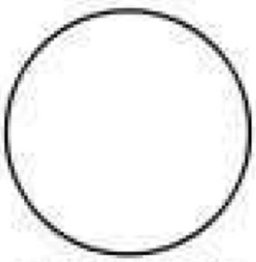
We are a team of Math & AI researchers at Caltech, focused on developing AI systems that can tackle hard research-level math problems. Solving challenging mathematical tasks --- such as proving or disproving long-standing conjectures, or establishing difficult theorems --- often requires discovering intricate, multi-step solutions. Our mission is to use these hard mathematical problems as *environments* to design new AI algorithms and architectures that can identify rare solutions carrying disproportionately high rewards.

Career opportunities

Postdoctoral positions

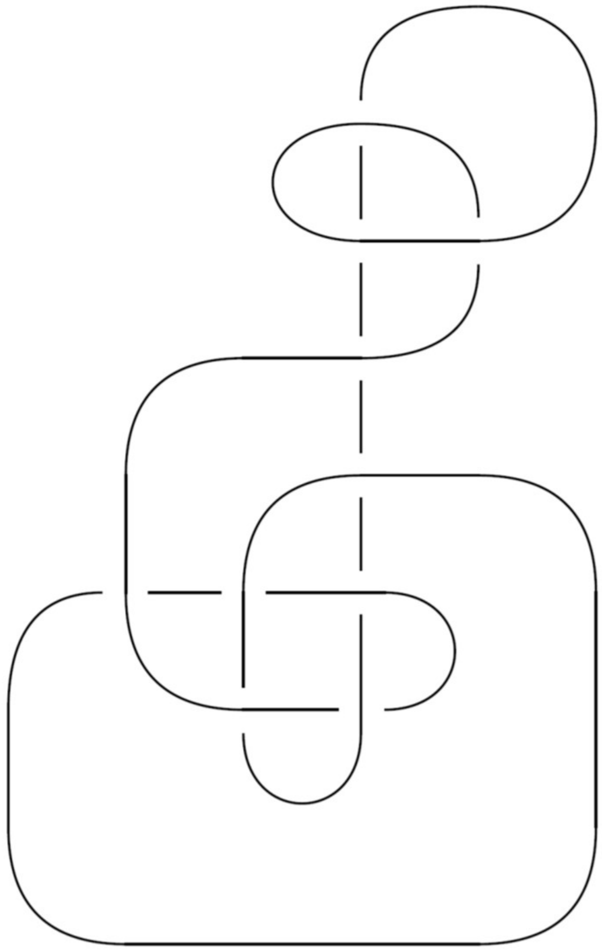
Caltech

1200 East California Boulevard
Pasadena, California 91125

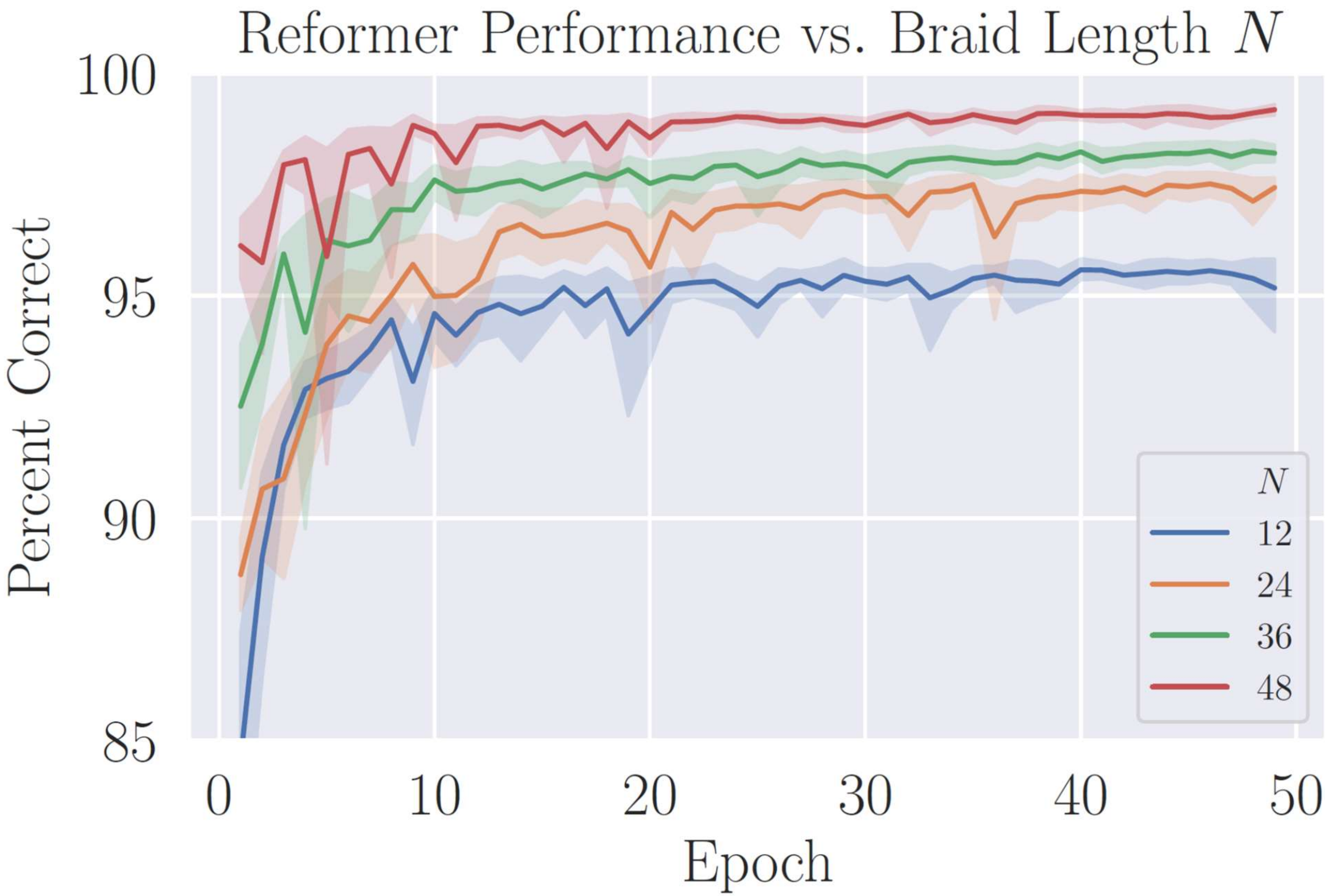


Unknot

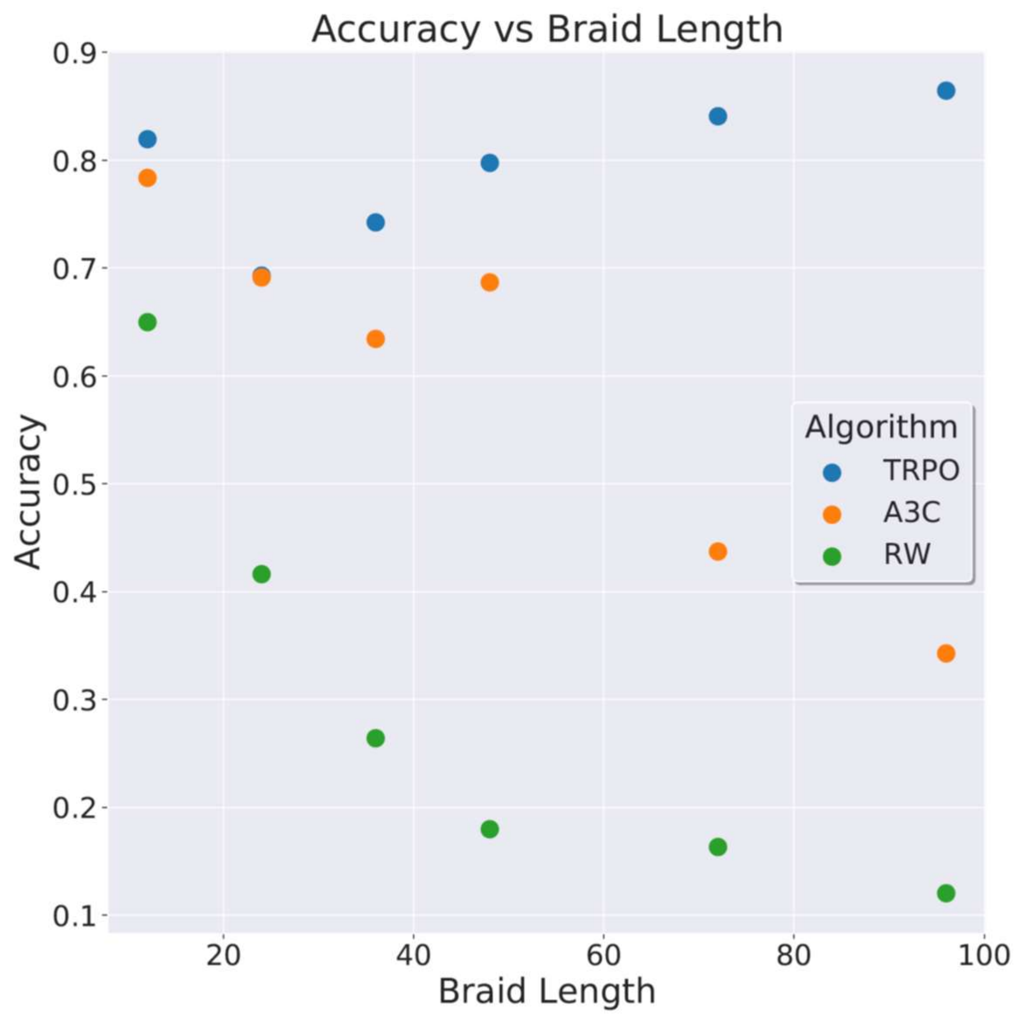
= ?



10 crossings



Reformer performance on UNKNOT as function of braid length. Performance increases with N.



Fraction of unknots whose braid words could be reduced to the empty braid word as a function of initial braid word length.